THE FULL MOTION SYSTEM FOR CD-I

Jan van der Meer
Philips Consumer Electronics
Eindhoven, The Netherlands

ABSTRACT

Compact Disc Interactive (CD-I) is the new multimedia system for home entertainment, education, training, publishing and other applications that combines audio, video, text, graphics, animation and interactivity. To extend CD-I with the capability to play moving natural pictures on full screen with associated audio of Compact Disc quality, the Full Motion system is defined. To play Full Motion sequences from CD-I disc requires compression of the audiovisual information to the CD-I bitrate. The Full Motion system applies a compression method based on the MPEG standard. This paper describes the features of the CD-I Full Motion system, application of the MPEG standard for Full Motion and the architecture of a Full Motion CD-I player.

1. INTRODUCTION

After an initial introduction on the professional and institutional market, CD-I has been introduced on the consumer market in the USA in October 1991, followed by Japan and United Kingdom in April 1992; the rest of Europe followed in September and October 1992. CD-I combines audio, video, text, graphics, animation and interactivity. The capability to play moving natural pictures is limited to partial screen (e.g. 1/9 of the screen) and low picture rates (e.g. 15 Hz). To improve this performance the Full Motion system is defined.

The Full Motion system introduces in CD-I the capability to play from one CD-I disc up to 72 minutes of moving natural pictures on full screen with audio on Compact Disc quality. Each Full Motion CD-I disc can be played on 50 Hz and a 60 Hz players. The Full Motion system offers a world standard with full compatibility between the 50 Hz and 60 Hz world.

The Full Motion system is defined as an extension of the CD-I Specification [1], without requiring any modification of this specification. As a consequence, Full Motion is a full compatible extension of CD-I. Each CD-I disc without Full Motion plays on CD-I players with Full Motion. Furthermore, each Full Motion CD-I disc "plays" also on CD-I players without Full Motion, although the Full Motion functions will not be available on such players.

CD-I discs are produced at authoring. The Full Motion system provides a variety of features to design Full Motion CD-I discs. An important feature at authoring is flexibility to optimize the discs for an application. For example, a trade-off can be made between audio quality and the number of audio streams. At a bitrate of 192 kbit/sec e.g. one audio stream in stereo of CD quality can be recorded or six audio streams in mono of speech quality. At Full Motion CD-I players a number of features is available to play back Full Motion CD-I discs. Section 2 of this paper presents the features provided by the Full Motion system at authoring and at the Full Motion CD-I player.

A typical Full Motion CD-I disc is expected to allocate 1.2 Mbit/sec for video on full screen, and 0.2 Mbit/sec for audio in stereo of CD quality; the total bitrate from disc equals 1.4 Mbit/sec. To compress the audiovisual information to these bitrates and to store the coded audio and video data on CD-I discs, the MPEG standard [2] is applied. The MPEG standard is defined by ISO/IEC as a generic standard, i.e. as an application independent standard. The application of this generic standard for Full Motion is discussed in section 3 of this paper.

In section 4 of this paper the architecture of the Full Motion CD-I player is presented. The Full Motion system is defined such, that 'Base Case' players, i.e. players without any extension, can be upgraded easily with Full Motion. In the architecture of the Full Motion CD-I player, Full Motion is an easy add-on function. This approach made it possible to prepare Base Case CD-I players for Full Motion. A consumer can upgrade a prepared player with Full Motion simply by installing a Full Motion cartridge (see figure 1).

Manuscript received October 5, 1992

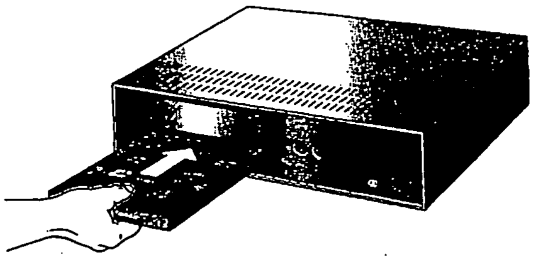

Figure 1 Installing a Full Motion cartridge in a Base Case CD-I player prepared for Full Motion.

2. FULL MOTION FEATURES

The Full Motion System provides a variety of features both at authoring and at the Full Motion CD-I Player. At authoring features are available to support the design of Full Motion CD-I discs. Flexibility to optimize discs to the requirements of the application is an important feature at authoring. Features available at a Full Motion CD-I Player support the user to play Full Motion CD-I Discs. These features include play back modes, presentation of played audiovisual information and full compatibility of discs between the 60 Hz and 50 Hz world. Beyond the features directly related to Full Motion, two additional features are provided. These features address handling of still pictures and use by applications of the available memory in a Full Motion Decoder, when this memory is not needed for decoding Full Motion data.

2.1 Features at authoring

Full Motion CD-I discs are produced at authoring. The CD-I application is specified and developed. The Full Motion audiovisual material is compressed in compliance with the MPEG standard. The resulting coded data are merged with the other CD-I data to create the Full Motion CD-I disc image.

After compressing an audio or video input signal, the coded data is stored in an MPEG audio or MPEG video stream. On a CD-I disc multiple MPEG audio and MPEG video streams can be recorded in parallel, e.g. for applications requiring audio in several languages. At the Full Motion CD-I Player, only one MPEG audio stream and one MPEG video stream can be played simultaneously. At authoring trade-offs can be made between the number of streams, the bitrate for each stream, the audio or picture quality of each stream and, in case of video, the picture size.

Audio can be coded in stereo and in mono. All fixed bitrates allowed by the MPEG standard can be used. The lowest bitrate for mono of 32 kbit/sec can be used for speech and other audio not requiring a very high quality. At a bitrate of 192 kbit/sec in stereo an audio quality can be achieved which is subjectively equal to the Digital Audio quality from Compact Disc.

The Full Motion system codes non-interlaced moving video. To support source material in 60 Hz and 50 Hz, several temporal formats can be used. Picture rates of (approximately) 24 Hz and 30 Hz support film and video based source material in 60 Hz, while source material in 50 Hz is supported by a picture rate of 25 Hz. The spatial format is a rectangle with a horizontal size in pixels and a vertical size in lines. The maximum picture area is limited; in case of 30 Hz to 352 pixels by 240 lines, and in case of 25 Hz to 352 pixels by 288 lines. However, the picture area can be used in a flexible way. Full screen pictures can be coded, but also wide and low pictures as well as high and narrow pictures (for an example see figure 2).

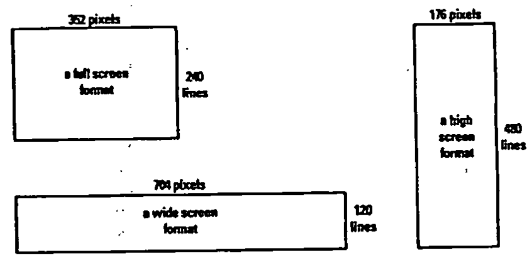

Figure 2 Example of spatial picture formats.

The bitrate to code video is a parameter. In case of a play back directly from a Full Motion CD-I Disc, the maximum bitrate is about 1.4 Mbit/sec. After storage of the sequence by the application, play back is also possible from memory. In that case a maximum bitrate of approximately 5 Mbit/sec is allowed e.g. for playing video with a very high picture quality.

Each Full Motion CD-I Disc can be played on each Full Motion CD-I Player, independently of the 60 Hz or 50 Hz display rate of the player. The Full Motion decoder reconstructs pictures at the picture rate they are coded, i.e. at about 24 Hz, 30 Hz or 25 Hz. The decoder is to produce a video output at a display rate of 60 Hz or 50 Hz. The Full Motion decoder therefore needs to apply a frame rate conversion from the coded picture rate to the required display rate of 60 Hz or 50 Hz. The conversion is applied in temporal direction without any impact on picture size in terms of number of pixels and lines. Due to this conversion, some

aspect ratio distortion may occur. At authoring a pixel aspect ratio can be applied which is optimum for either a 60 Hz or a 50 Hz display, but which results in a noticeable distortion of the aspect ratio when the pictures are displayed at the other display rate. For compatibility reasons therefore an intermediate pixel aspect ratio is recommended which results in a minor distortion of the aspect ratio at both 60 Hz and 50 Hz displays; without reference to the original, this distortion is not noticeable in both cases.

For storage on a Full Motion CD-I disc, the audio and video material is compressed to the CD-I bitrate. A typical encoding environment is depicted in figure 3. The video signal from the source is filtered and sub-sampled to the SIF format [2] by a pre-processor. The SIF pictures are input to a video encoder. The audio signal from the source is provided to the audio encoder; in case of very low bitrates (speech quality) the audio may be pre-filtered first. In figure 3 some typical values are indicated for bitrates and bitrate reduction (BRR). To obtain highest picture quality, it is essential that the SIF pictures at the input of the video encoder are natural pictures of high quality, without spatial and temporal distortions. In general the pre-processor will apply therefore techniques to reduce e.g. noise and alias.

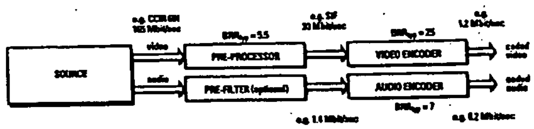

Figure 3 Typical audio and video encoding environment

The Full Motion system specifies the format of the coded bitstreams at the output of the encoders. The techniques to be applied for pre-processing and encoding techniques are not specified; they are at the discretion of the authoring tool. A major feature of the Full Motion system is therefore that each improvement of the encoding strategy can be applied at encoding. A continuous growth path to improve picture and audio quality is therefore expected, without any need to modify Full Motion CD-I players.

Applications can apply a variety of encoding tools. Audio encoding tools are available already for operation in real time. Currently video encoders are available for operation in non-real time. The speed of non-real time video encoders is increasing rapidly, due to speed optimization of the algorithms as well as due to faster computer based encoding platforms. Video encoders operating in real time are expected by the end of 1992.

Especially flexibility is available in coding video. To code a video sequence, the MPEG standard allows to use three types of pictures : Intra-pictures (I), Predicted pictures (P) and Interpolated pictures (bidirectional prediction : B). Intra-pictures provide entry points for random access, but only with moderate compression. Predicted pictures are coded with reference to a past picture (I or P picture). Bidirectional (B) pictures provide the highest amount of compression, but require both a past and a future reference picture (I or P picture) for prediction.

An example is given in figure 4; in this example the I and P picture are followed by two B pictures. The number of B pictures between two successive I or P pictures may vary. Coding of a B picture requires availability of the past and future reference picture. Each future reference picture is therefore coded prior to the B picture(s) using this reference picture. See figure 4. For each video sequence, the encoder decides which picture types are applied. Also the amount of bits to code each picture is at the discretion of the encoder.

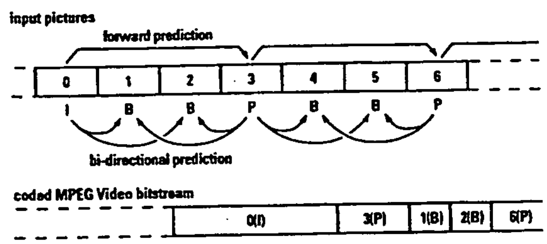

Figure 4 MPEG Video encoding

At the decoder, I pictures are reconstructed without using any information from other pictures. At these type of pictures play back can start. They are referred to as entry point pictures. The distance between entry point pictures is flexible. Applications can specify entry point pictures in a video sequence prior to encoding the sequence.

2.2 Features of Full Motion CD-I Player

Each Full Motion CD-I disc contains an application program. This application program controls play back of the disc on the Full Motion CD-I player, in interaction with the user. Part of such control is to decide which MPEG video stream and which MPEG audio stream are to be played, if any. The application can control separately decoding of MPEG audio data and decoding of MPEG video data. The Full Motion system allows for play back directly from disc, as well as from memory. In case of a play from memory, the data to be played is loaded in memory by the application. Play from disc and play from memory are mutually exclusive functions for audio as well as for video. The Full Motion system supports loops from memory, by means of which MPEG data stored in memory can be played continuously.

At the Full Motion CD-I player, MPEG audio can be played simultaneously with ADPCM audio. To play back MPEG audio, the play forward mode and the mute mode are available. In the play forward mode, MPEG audio is played back at normal speed forward. A play forward can be paused and continued. In the mute mode the audio output from the Full Motion decoder is muted. The audio signals at the output of the Full Motion decoder are provided to the external audio input of the Base Case audio section (see figure 5). The left and right MPEG audio signals are mixed with the left and right audio output signals from the ADPCM decoder.

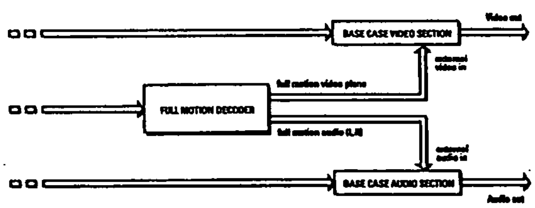

Figure 5 Mixing of audio and video from Full Motion decoder with Base Case audio and video

Coded MPEG Audio data is a concatenation of coded audio frames. An audio frame represents a number of audio samples. The decoder can reconstruct each audio frame without using information from other audio frames. Play back of MPEG Audio can therefore start at each audio frame.

On the play back modes of MPEG video there is some impact from the predictive video coding techniques applied in the MPEG standard. Play back can start at entry point pictures (I pictures). Entry point pictures are defined during encoding. The address of entry point pictures on a disc is known by the application. For reconstruction of P and B pictures, previously reconstructed reference pictures are required; therefore at these picture types play back cannot start.

At the decoder pictures are reconstructed in the same order as they are included in the bitstream. After reconstruction pictures need re-ordering. See figure 6. The re-ordered pictures are to be displayed at a frame rate of 50 Hz or 60 Hz, depending on the player type used for play back. The decoder applies a frame rate conversion to the required display rate. See figure 6.

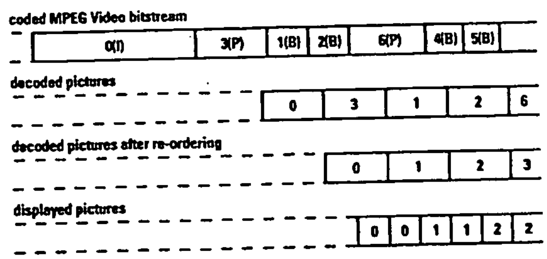

Figure 6 MPEG Video decoding

To play back MPEG video, five modes are available in the Full Motion system : play forward, slow motion forward, freeze, single picture forward and scan. In the play forward mode MPEG video is played back at normal speed forward.

In the slow motion forward mode MPEG video is played forward in slow motion. The application can select the slow motion speed from seven available speeds, defined by the formula :

slow motion speed = (1/m) * normal speed,
where m = 2, 3, 4, ..., 8.

See example in figure 7. A play forward and a slow motion can be paused at each picture and can be continued at the paused picture.

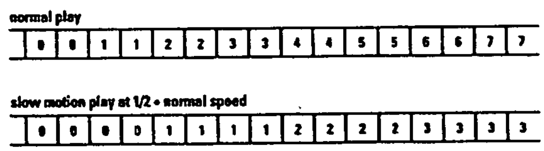

Figure 7 Displayed pictures at normal play and at slow motion

The freeze mode can be entered from the play mode at each picture. At entering the freeze mode, the currently displayed picture is frozen and the decoding stops, but the MPEG video stream continues. From the freeze mode, the play mode can be continued. In that case play back will continue at the first entry point picture (I picture) encountered in the MPEG video bitstream. For an example see figure 8.

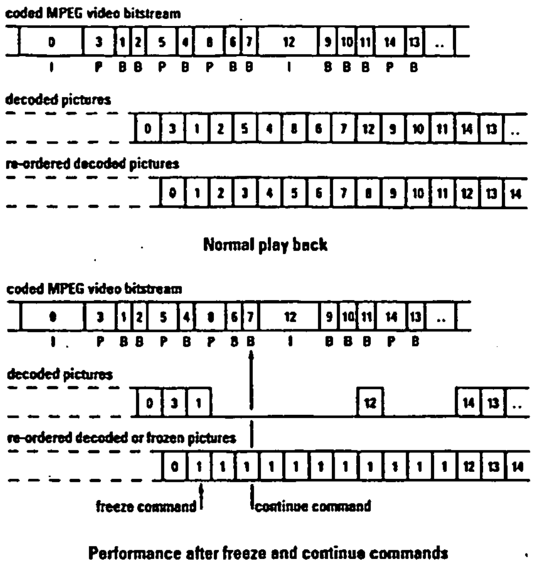

Figure 8 Normal play back and performance after freeze and continue commands; in this example the freeze command is issued during display of picture no 1, while the continue command is issued prior to encountering I picture no 12 in the MPEG video bitstream

In the single picture forward mode, pictures are played back one by one. By giving another single picture forward command, the next picture is displayed. The application can control the length of the display period of each picture by controlling the time interval between successive commands.

During the scan mode entry point pictures are seeked, reconstructed and displayed as a still picture for a nominal display period, during which the next entry point picture is seeked and decoded. The length of the nominal display period is controlled by the application. The application may scan entry point pictures in forward, reverse or random direction.

After decoding the MPEG video data, the Full Motion decoder displays reconstructed MPEG video pictures or a part thereof in the full motion video plane, available at the output of the decoder. The full motion video plane has a horizontal resolution of 384 or 360 pixels on a line, and is fully synchronized with the other planes, i.e. the base case video planes. The width of all planes is 384 or 360 pixels of normal resolution and the height of all planes is 240 or 280 lines. The field rate of the full motion video plane can be 60 Hz or 50 Hz, either interlaced or non-interlaced, exactly following the vertical timing of the base case video planes.

The application controls the part of the reconstructed picture to be displayed by means of a display window. The display window is a rectangle within the reconstructed picture; the reconstructed picture area within the display window is displayed in the full motion video plane. See figure 9. The reconstructed picture or the part thereof which is displayed may be smaller than the spatial size of the full motion video plane. In areas of the full motion video plane not covered by the displayed part of a reconstructed picture, a background color is displayed. The application can program one background color in RGB for the entire full motion video plane.

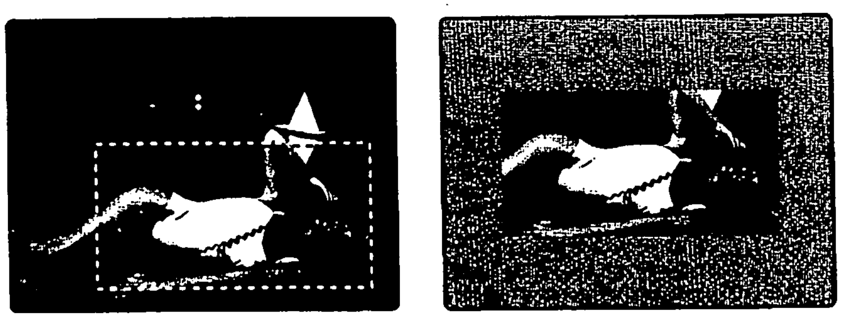

Figure 9 Display window from reconstructed picture displayed within full motion video plane

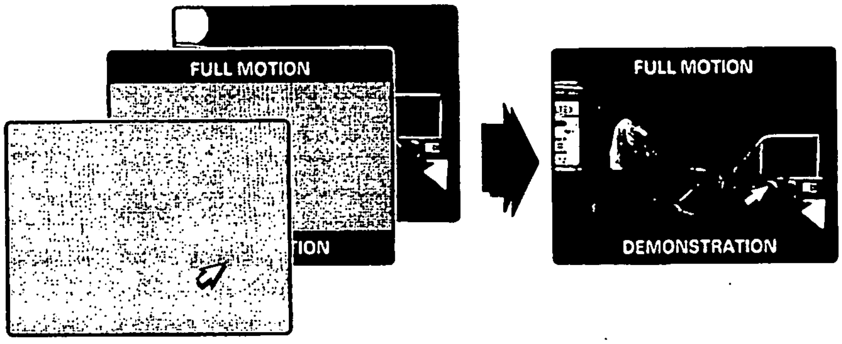

Figure 10 Overlay of the cursor plane, a base case (e.g. RGB) plane and the full motion video plane

The application can control the size and the position of the display window by means of the following parameters :
- height and width of window;
- position of window within reconstructed picture;
- position of window within full motion video plane.

By means of above parameters a variety of display functions can be realized. For example the window can be opened or closed slowly by modifying the size of the window each field period. Also scrolling within the reconstructed picture is possible, as well as scrolling within the full motion video plane. Another function available to applications is to blank the window; in that case the contents of the display window becomes 'black', i.e. R,G,B = 16,16,16.

At the output of the Full Motion decoder the full motion video plane is available. The full motion video plane is provided to the external video input of the Base Case video section (see also figure 5). By controlling this input, the application can replace the backdrop plane by the full motion video plane. When enabled, the full motion video plane will be visible in picture areas where transparency is defined for all enabled base case video planes, including the cursor (see figure 10). Note that an application can apply pixel accuracy in defining transparency.

2.3 Still Pictures

A still picture is a video sequence consisting of one picture. Non-interlaced still pictures with a picture area of up to 352 pixels by 288 lines can therefore be coded as a normal full motion sequence. For such still pictures the previous sections 2.1 and 2.2 are applicable.

The technique to compress still pictures, as provided by the MPEG standard, is more efficient than the methods available in a Base Case player. For a typical still picture, MPEG compression is approximately four times more efficient as DYUV compression (DPCM based) in the Base Case player. In a special mode, the Full Motion system therefore also supports MPEG based coding of interlaced still pictures.

For coding of interlaced still pictures the maximum picture area is 384 pixels by 576 lines. Flexibility is available to exploit this area for coding wide and low pictures as well as high and narrow pictures. After reconstruction, interlaced still pictures are displayed in the full motion video plane, with both the full motion video plane and the base case video planes in the interlace mode. The part of the still picture to be displayed is selected by means of a display window, as for non-interlaced pictures. The control of the full motion video plane, the control of the display window and the mixing with base case video planes is the same as for non-interlaced video; see section 2.2.

Default one reconstructed interlaced still picture can be stored. An extension mechanism is defined for Full Motion decoders with larger memories to store multiple reconstructed still pictures. The application can read the number of interlaced still picture which can be stored in the Full Motion decoder from a Device Status Descriptor (DSD). A DSD contains characteristics of a specific device in a player.

2.4 Memory extension

The Full Motion decoder in a CD-I player contains a memory bank of at least 4 Mbit, a.o. to handle MPEG video decoding and re-ordering as well as frame rate conversion. When the memory is not used for Full Motion video decoding, the memory can be used to extend the amount of system memory available to applications. In that case the Full Motion decoder serves as memory controller, such that the memory is transparently accessible by the application. The minimum size of the memory bank is 4Mbit. In the general case, the size of the memory bank applied in a Full Motion CD-I player is found from the formula :

Size memory bank = (B * 1) Mbit,

where B >= 4; B is an integer value. The application can find the size of the applied memory bank from a Device Status Descriptor (DSD), where the applied value of B is stored.

3. APPLICATION OF THE MPEG STANDARD FOR FULL MOTION

The MPEG standard is a generic standard for audiovisual coding, i.e. an application independent standard. The MPEG standard has been defined by working group WG11 of subcommittee SC29 of the joint technical committee JTC1 of the standardization bodies ISO and IEC. In this working group, the Moving Picture Experts Group (MPEG), experts from many countries, representing a.o. the computer industry, the VLSI industry, the telecommunication industry, and the consumer electronics industry, contributed to develop the MPEG standard.

The three companies collaborating in the Full Motion System for CD-I, Matsushita Electric Industry, Sony Corporation and Philips Consumer Electronics, decided to apply the MPEG standard for this system. The techniques developed by the MPEG Committee will enable many applications requiring digitally compressed video and audio. The storage media targeted by MPEG include, next to Compact Disc, also DAT and computer disks and it is expected that MPEG based technologies will eventually be used in a variety of communication channels, such as ISDN and local area networks and even in broadcasting applications.

The MPEG standard specifies an MPEG stream as a generic interchange format. To store an MPEG stream on a medium, a medium specific layer is required (see figure 11). In case of the Full Motion system, the medium is the CD-I disc or memory and the medium specific layer is the disc format.

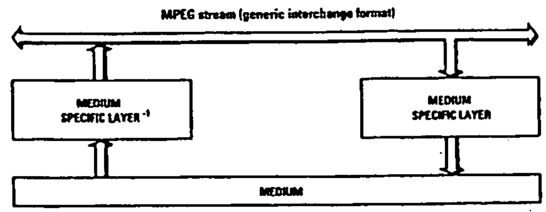

Figure 11 Storage of an MPEG stream on a medium.

The MPEG standard consists of three parts. In part 1 the system layer of the standard is specified, including the multiplex structure to combine MPEG audio and MPEG video data; furthermore timing information needed for synchronized play back is specified in part 1. [3] In part 2 of the MPEG standard the video coding layer is specified. [4][5] Part 3 specifies the audio coding layer.

A general MPEG encoding system is given in figure 12. Audio and video sources are providing inputs to audio and video encoders. The MPEG Audio and MPEG Video streams generated by the encoders are multiplexed into one MPEG stream. In one MPEG stream up to sixteen MPEG Video streams and up to thirty-two MPEG Audio streams can be multiplexed. Furthermore a System Time Clock is available with a nominal value of 90 kHz. The System Time clock is used for synchronization.

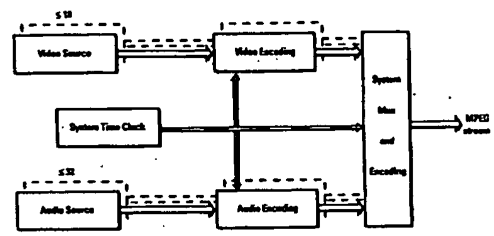

Figure 12 General MPEG encoding system

The MPEG multiplex structure consists of packs and packets. Each pack consists of a pack header and a number of packets. Each packet consists of a packet header and a data field; the data field contains data from one of the streams in the multiplex. For an example see figure 13.

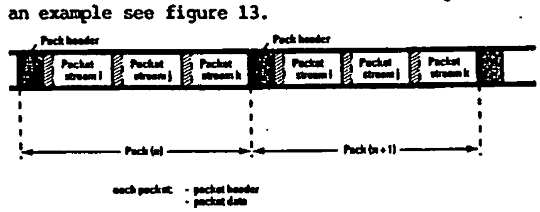

Figure 13 Example of MPEG multiplex structure with three streams (i, j and k).

For synchronization reasons, values of the System Time Clock are encoded in packet headers and pack headers. In packet headers Time Stamps (TS) are encoded. A Time Stamp indicates the value of the System Time Clock at the instant that a picture or audio frame enters the video or audio encoder. See figure 14. In the pack header the System Clock Reference (SCR) is encoded, indicating the value of the System Time Clock at the instant the SCR field exits the encoding system as part of the produced MPEG stream. At decoding, the inverse process is followed. From the SCR values in the pack headers, the System Time Clock is regenerated. Decoded pictures and audio frames are output of the decoders at the instants indicated by the Time Stamps. As a result, the mutual timing relations between audio and video at the output of the decoder are the same as at the input of the encoder.

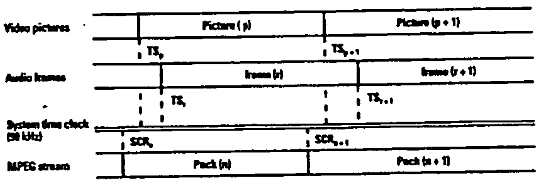

Figure 14 Synchronization between audio and video in MPEG, using the 90 kHz System Time Clock.

In the MPEG standard a number of parameters are defined; examples are the picture rate, the picture size, the audio sampling frequency and the bitrate. To make the MPEG standard suitable for many applications, these parameters usually have a large range. The picture size e.g. can be up to 4096 pixels by 4096 lines.

To define the parameter range covered, an application has to specify its 'MPEG profile'. The MPEG profile also defines the specific constraints from the application on the MPEG syntax and semantics. In the MPEG standard an MPEG profile has been defined already for target applications. The Full Motion system adopts the profile of CSPS (Constrained System Parameter Stream) as defined in the MPEG standard for the system layer, while the profile for the video coding layer of constrained parameters is supported. In the Full Motion system some video parameters are allowed to go beyond the constrained parameter range for video.

Next to its MPEG profile, an application has to specify the application specific characteristics. Examples are the medium specific layer (see figure 14) and how to use the MPEG stream in the application. For the Full Motion system the latter is summarized in section 2 of this paper. The medium specific layer for Full Motion CD-I discs is summarized in section 3.1. The MPEG profile applied by the Full Motion system is summarized in sections 3.2, 3.3 and 3.4 for the audio, video and systems parts of the MPEG standard. In section 3.5 synchronized playback of Full Motion sequences is described.

3.1 The medium specific layer for storage on CD-I disc

On a CD-I disc data is stored in sequentially recorded units called sectors. At a player CD-I discs are read at a constant rate of 75 sectors per second. Each CD-I sector consists of a header, a subheader and a datafield. The header and subheader provide a.o. information on the data type, the form, and coding information. The data type indicates the type of data stored in the sector; video sectors, audio sectors and data sectors are distinguished. The form (which may be 1 or 2) indicates whether error correction code is included in the datafield. Form 1 sectors contain 2048 Bytes of user data and 280 Bytes of error correction code. Form 2 sectors contain fully 2324 Bytes of user data without any error correction code. The coding information indicates the applied coding format of the sector. For example video sectors may contain DYUV, CLUT or RGB data and audio sectors may contain ADPCM data on compression level A, B or C.

To store Full Motion data on CD-I discs, the coding format options are extended. Two new coding formats are defined for video sectors : 'MPEG video' and 'MPEG still picture' for storage of MPEG coded non-interlaced moving picture data and interlaced still picture data. One new coding format is defined for audio sectors : 'MPEG audio' for storage of coded MPEG audio data. Each MPEG video sector, MPEG still picture sector and MPEG audio sector is a form 2 sector with a usable datafield of 2324 Bytes. Note that the Full Motion application program is stored in a normal data sector.

To store Full Motion data in sectors, the system part of the MPEG standard is applied. The multiplex structure provided by the system part is applied for a separate multiplex of MPEG video streams, MPEG still picture streams and MPEG audio streams. As a result three type of ISO11172 streams can be distinguished:

a) In this type of ISO 11172 stream MPEG video streams are multiplexed. In each ISO11172 stream of this type one or more (up to 16) MPEG video streams may be multiplexed. This type of ISO11172 streams is stored in MPEG video sectors.
b) In this type of ISO 11172 stream MPEG still picture streams are multiplexed. Each ISO11172 stream of this type may be a multiplex of one or more (up to 16) MPEG still picture streams. This type of ISO11172 streams is stored in MPEG still picture sectors.
c) In this type of ISO 11172 stream MPEG audio streams are multiplexed. In each ISO11172 stream of this type one or more (up to 32) MPEG audio streams may be multiplexed. This type of ISO11172 streams is stored in MPEG audio sectors.

Each ISO11172 stream is stored in an integer number of sectors.

IEEE Transactions on Consumer Electronics, Vol. 38, No. 4, NOVEMBER 1992

3.2 MPEG audio profile

To code audio, the Full Motion system applies layer I and layer II of the audio part of the MPEG standard. Audio can be coded in stereo, in joint stereo (i.e. intensity stereo), in dual channel or in single channel. In case of intensity stereo, the redundancy between both stereo channels is exploited to improve the coding efficiency. All bitrates allowed by MPEG for layer I and layer II are supported; see figure 15. The applied audio sampling frequency is 44.1 kHz. For simple integration of the MPEG Audio decoding functions with Base Case decoding functions, other sampling frequencies are not allowed in the Full Motion system. The protection bit may be used to add redundancy to the MPEG audio data. Two options may be used for emphasis : no emphasis and 50/15 μsec emphasis. Other options, including CCITT J.17 emphasis are not allowed.

Layer I			Layer II		
bitrate	mono	stereo	bitrate	mono	stereo
32	yes	yes	32	yes	no
64	yes	yes	48	yes	no
96	yes	yes	56	yes	no
128	yes	yes	64	yes	yes
160	yes	yes	80	yes	no
192	yes	yes	96	yes	yes
224	yes	yes	112	yes	yes
256	yes	yes	128	yes	yes
288	yes	yes	160	yes	yes
320	yes	yes	192	yes	yes
352	yes	yes	224	no	yes
384	yes	yes	256	no	yes
416	yes	yes	320	no	yes
448	yes	yes	384	no	yes

Figure 15 Available audio bitrate options in kbit/sec
Note : stereo refers to dual channel mode, independent stereo mode or intensity stereo mode; mono refers to single channel mode.

3.3 MPEG video profile

The Full Motion system supports the constrained video parameters, defined in the video part of the MPEG standard, but some parameters have a larger range in the Full Motion system. The Full Motion system allows for a maximum bitrate for video of about 5 Mbit/sec. This bitrate exceeds the 1.856 Mbit/sec limit defined in the MPEG standard for the constrained video parameters. Furthermore, for coded interlaced still pictures the maximum picture size exceeds the picture size limit for the constrained video parameters. Finally, in the Full Motion system video coding with a variable bitrate is allowed.

The following summarizes the bounds specified in the video part of the MPEG standard for the constrained parameters. See the MPEG standard for the definition of parameters. The maximum picture width is 768 pixels, and the maximum height is 576 lines. The maximum picture area is 396 macroblocks. A macroblock is a rectangular block of 16*16 luminance pixels and corresponding chrominance pixels. The maximum number of macroblocks per second equals 396*25. The constrained parameters allow for picture rates of 23.976 Hz, 24 Hz, 25 Hz, 29.97 Hz and 30 Hz. The maximum size of the VBV buffer is 40 kByte, and the range of motion vectors on half pixel accuracy is bounded to approximately +/- 64 pixels. For further details see part 2 of the MPEG standard.

An MPEG video data stream may be a concatenation of sequences with different parameters. For example the picture rate or the picture size may change. The Full Motion system applies no constraints on the pixel aspect ratio. Applications should take into account however that Full Motion decoders do not compensate for aspect ratio distortion. Three aspect ratios options are recommended for practical use : one optimum ratio for 60 Hz displays, one optimum ratio for 50 Hz displays, and one intermediate ratio suitable for both 60 Hz and 50 Hz displays.

3.4 MPEG system profile

The system part of the MPEG standard applies a multiplex structure consisting of packs and packets. See also figure 13. In each MPEG video sector, MPEG still picture sector and MPEG audio sector one pack is stored. Each pack consists of a pack header and one or more packets. A packet consists of a packet header followed by data from one stream.

The System Clock Reference (SCR) parameter in the pack header indicates, in units of the System Time Clock, the nominal arrival time of the pack at the input of the Full Motion decoder (see also figure 14). At play back, the decoder uses the SCR parameter to control the System Time Clock. For example at the start of a play back, the decoder System Time Clock can be initialized to the value equal to the SCR value encoded in the header of the first pack entering the decoder.

In the Full Motion system the System Time Clock is locked to the rate of 75 Hz at which sectors are read from disc. As a consequence, in case of a play from disc, the difference between encoded SCR values in sectors from the same ISO11172 stream is a multiple of 1200 (1200*75 = 90000). Full motion data can be played from memory with sector rates of 1, 2, 3 or 4 times the nominal sector rate. In general therefore the SCR difference described above is a multiple of 1200, 600, 400 or 300.

In the Full Motion system the System Time Clock is locked furthermore to both the picture rate in MPEG video streams and MPEG still picture streams, and to the audio sampling frequency of 44.1 kHz. The locking to the picture rate and the audio sampling frequency is ensured at encoding.

Each ISO11172 stream within the Full Motion system is a CSPS stream (Constrained System Parameter Stream; see system part of MPEG standard).

3.5 Playback synchronization

Synchronization between MPEG audio and MPEG video is ensured implicitly by the MPEG standard. The decoder regenerates the System Time Clock and presents decoded pictures and audio frames based on Time Stamps. See also figure 14. In the Full Motion system, the MPEG audio decoder and the MPEG video decoder do not share the same System Time Clock. Both decoders have their own System Time Clock. Because both clocks are locked to the 75 Hz sector delivery rate, their mutual offset is constant during a play. The mutual offset to be applied is known by the application. At playback, synchronization between MPEG audio and MPEG video is achieved, when the required mutual offset is applied between both System Time Clocks.

The MPEG standard assumes instantaneous decoding without any processing delay. In practice processing delays occurs in MPEG audio and MPEG video decoders. Each processing delay consists of a variable and a fixed component. The variable component is to be within 40 msec for each MPEG audio or MPEG video decoder in a Full Motion CD-I player. Within a player, the fixed component is to be constant for each MPEG audio or MPEG video decoder. The maximum synchronization error in the presentation of MPEG video and MPEG audio is therefore +/- 40 msec. The value of the constant component of the MPEG processing delay is player dependent and stored in a DSD.

Synchronization between a Full Motion play and plays of ADPCM audio and Base Case video is the responsibility of the application. To compensate for the processing delay of MPEG audio and MPEG video data, applications should delay the ADPCM audio play and the Base Case video play. The required value for such compensating delay can be found from the constant component of the MPEG processing delay stored in a DSD. To make synchronization possible between application events and Full Motion events, the Full Motion decoder informs the application, upon request, about Full Motion events. Examples are 'start of display of next picture' and 'presentation starts of next audio frame'.

4. ARCHITECTURE OF FULL MOTION CD-I PLAYER

A CD-I Base Case player is extended with the capability to play Full Motion CD-I discs by implementing a Full Motion decoder in the player. The interface between the Full Motion decoder and the Base Case player is straight forward (see figure 16). The Full Motion decoder is connected to and controlled via the CPU bus. The transfer of Full Motion data from disc or memory to the decoder is also via the CPU bus. The audio and video outputs from the Full Motion decoder are provided to the external audio and video inputs of the Base Case audio and video sections. For video synchronization the pixel clock as well as horizontal and vertical synchronization signals from the Base Case video section are provided to the Full Motion decoder. To lock the decoders System Time Clocks to the sector rate, a CD data clock is provided to the Full Motion decoder; this clock can be e.g. the master clock to which data delivery by the CD drive is locked. The frequency of such clock is player dependent.

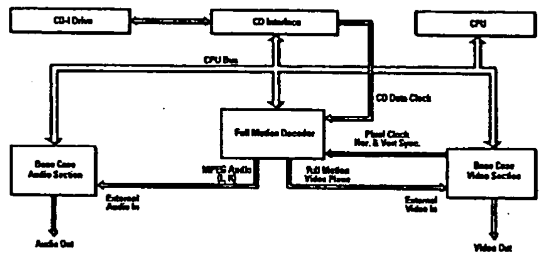

Figure 16 Full Motion decoder in a CD-I player

4.1 Structure of the Full Motion decoder

The Full Motion decoder basically consists of two functions, an MPEG video decoding function and an MPEG audio decoding function. Both MPEG decoding functions are in practice realized in a combination of hardware and software; which parts are realized in hardware and which in software is implementation dependent. The MPEG video processing hardware may consist of an MPEG video decoder IC with an external memory bank of at least 4Mbit. The memory bank can be used for multiple purposes during MPEG video decoding :

a) for buffering of coded input data,
b) for storage of reference pictures at decoding,
c) for re-ordering of decoded pictures, and
d) as display memory at frame rate conversion.

Next to the input buffer of 46 kByte, as defined by the MPEG standard, in total slightly more than three pictures of maximum size can be stored in 4 Mbit.

The MPEG audio decoder hardware may consist of an MPEG audio decoder IC. The MPEG video and audio decoding software are the MPEG video driver and the MPEG audio driver. Both drivers are hardware dependent. The application controls the Full Motion decoder by communicating with the MPEG video driver and the MPEG audio driver. The application communicates with both drivers via the Full Motion manager (see figure 17). The Full Motion manager is part of CD-RTOS, the compact disc real time operating system applied in CD-I. For the Full Motion system, the specification of CD-RTOS is therefore extended with the Full Motion manager, the calls between the application and the Full Motion manager, and the calls between the Full Motion manager and both the MPEG video driver and the MPEG audio driver.

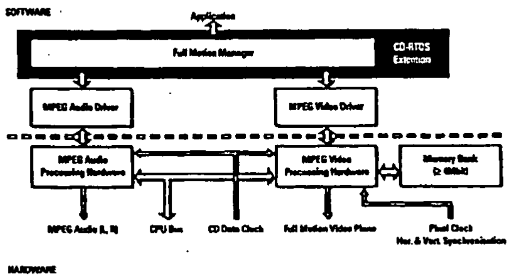

Figure 17 Structure of Full Motion decoder

5. CONCLUSION

In this paper the Full Motion system for CD-I is presented. The Full Motion System is a full compatible extension of the CD-I system. After compression, the audiovisual information is stored on CD-I disc at a data format which is fully MPEG compatible. In the architecture of a Full Motion CD-I player, Full Motion is an simple add-on function. The system provides a variety of features to produce and playback Full Motion CD-I discs. The major features are 50 Hz /60 Hz disc compatibility, a high futureproof flexibility, and the capability to play from one Full Motion CD-I disc up to 72 minutes of audiovisual material of high quality.

References

[1] Compact Disc Interactive Full Functional Specification
September 1990
Sony Corporation and NV Philips

[2] MPEG1 standard
Coding of moving pictures and associated audio for digital storage media up to about 1.5 Mbit/sec.
April 1992
Draft international standard ISO/IEC DIS 11172
ISO/IEC JTC1/SC29/WG11

[3] Le Gall, D.J. MPEG : A video compression standard for multimedia applications
Communications of the ACM
Volume 34, Number 4, April 1991

[4] Le Gall, D.J. The MPEG video compression algorithm
Signal Processing : Image Communication
Volume 4, Number 2, April 1992

[5] MacInnis, A.G. The MPEG systems coding specification
Signal Processing : Image Communication
Volume 4, Number 2, April 1992

Biography of Mr. Jan van der Meer

Jan van der Meer received his degree in electronic engineering from the Technical University of Twente in the Netherlands in 1979. He is currently Senior Engineer Video Signal Processing at Philips Consumer Electronics in Eindhoven, the Netherlands. He is responsable for the architecture and specification of the Full Motion system for CD-I. Previously he has been involved in the development of the HDMAC High Definition Television System.